BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 6: e94596
doi: 10.3897/biss.6.94596

OPEN ACCESS

Conference Abstract

Specimods: A web-based tool for producing Genbank submission files for sequenced museum specimens

Richard A. Levy‡, Jessie W. Berta-Thompson‡, Gary Olds‡, Andrew W. Wilson‡

‡ Denver Botanic Gardens, Denver, United States of America

Corresponding author: Richard A. Levy (ricklevy21@gmail.com)

Received: 08 Sep 2022 | Published: 09 Sep 2022

Citation: Levy RA, Berta-Thompson JW, Olds G, Wilson AW (2022) Specimods: A web-based tool for producing Genbank submission files for sequenced museum specimens. Biodiversity Information Science and Standards 6: e94596. https://doi.org/10.3897/biss.6.94596

Abstract

The expectation of published sequence data on Genbank and other sequence databases is not only a requirement for journal publication, but is incredibly valuable for continuing research in organismal evolution and ecology. The process of formatting and submitting sequence data from vouchered specimens is burdensome and tedious. Improperly formatted submissions lead to lengthy delays and frustrations on behalf of the submitters and reviewers. Specimods*[1] is a new tool that utilizes the Global Biodiversity Information Facility (GBIF) API to produce formatted source modifier files that will facilitate the batch upload of specimen sequences to Genbank, using the sequence submission platform. Once a template CSV is populated with the specimen catalog numbers, institution codes, SeqID, sequence descriptions, and the sequences, Specimods will pull specimen metadata from GBIF to produce two files: a FASTA file with a complete definition line (SeqID, organism name, sequence title) and the sequences as well as a source modifier file containing specimen metadata. Currently, Specimods supports Internal Transcribed Spacer (ITS) (1, 2, or both), SSUrRNA_18s, and LSUrRNA_28s sequences. The application*[2] uses an Express server and Node.js with data linked to a user account stored in a MySQL relational database. This presentation will demonstrate the tool using data from fungal specimens and show how submitting dozens of sequences with metadata can be accomplished in a matter of minutes.

© Levy R et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords

web tool, GBIF, API

Presenting author

Richard A. Levy

Presented at

TDWG 2022

Endnotes

*1 https://specimods.botanicgardens.org
*2 https://github.com/ricklevy21/fasta-gen